Time Optimal Control in Spin Systems

Navin Khaneja,[*] Roger Brockett,[†]   Steffen J. Glaser[‡]

February 1, 2008

Abstract

In this paper, we study the design of pulse sequences for NMR spectroscopy as a problem of time optimal control of the unitary propagator. Radio frequency pulses are used in coherent spectroscopy to implement a unitary transfer of state. Pulse sequences that accomplish a desired transfer should be as short as possible in order to minimize the effects of relaxation and to optimize the sensitivity of the experiments. Here, we give an analytical characterization of such time optimal pulse sequences applicable to coherence transfer experiments in multiple-spin systems. We have adopted a general mathematical formulation, and present many of our results in this setting, mindful of the fact that new structures in optimal pulse design are constantly arising. Moreover, the general proofs are no more difficult than the specific problems of current interest. From a general control theory perspective, the problems we want to study have the following character. Suppose we are given a controllable right invariant system on a compact Lie group, what is the minimum time required to steer the system from some initial point to a specified final point? In NMR spectroscopy and quantum computing, this translates to, what is the minimum time required to produce a unitary propagator? We also give an analytical characterization of maximum achievable transfer in a given time for the two-spin systems.

1   Introduction

Many areas of spectroscopic fields, such as nuclear magnetic resonance (NMR), electron magnetic resonance and optical spectroscopy rely on a limited set of control variables in order to create desired unitary transformations [5, 6, 7]. In NMR, unitary transformations are used to manipulate an ensemble of nuclear spins, e.g. to transfer coherence between coupled spins in multidimensional NMR-experiments [5] or to implement quantum-logic gates in NMR quantum computers [8]. However, the design of a sequence of radio-frequency pulses that generate a desired unitary operator is not trivial [9]. Such a pulse sequence should be as short as possible in order to minimize the effects of relaxation or decoherence that are always present. So far, no general approach was known to determine the minimum time for the implementation of a desired unitary transformation [6]. Here we give an analytical characterization of such time optimal pulse sequences related to coherence transfer experiments in multiple spin systems. We determine, for example, the best possible in-phase and

anti-phase [6, 10] coherence transfer achievable in a given time. We show that the optimal in-phase transfer sequences improve the transfer efficiency relative to the isotropic mixing sequences [11] and demonstrate the optimality of some previously known sequences.

During the last decade the questions of controllability of quantum systems have generated considerable interest [16, 17]. In particular, coherence or polarization transfer in pulsed coherent spectroscopy has received lot of attention [6, 9]. Algorithms for determining bounds quantifying the maximum possible efficiency of transfer between non-Hermitian operators have been determined [6]. There is utmost need for design strategies for pulse sequences that can achieve these bounds. From a control theory perspective, this is a constructive controllability problem [14]. At the same time it is desirable that the pulse sequences be as short as possible so as to minimize the relaxation effects. This naturally leads us to the problem of time optimal control, i.e. given that there exist controls that steer the system from a given initial to final state, we would like to determine controls that achieve the task in minimum possible time [17, 15].

In non-relativistic quantum mechanics, the time evolution of a quantum system is defined through the time-dependent Schrödinger equation

$$\dot{U}(t) = -iH(t)U(t), \quad U(0) = I,$$

where $H(t)$ and $U(t)$ are the Hamiltonian and the unitary displacement operators, respectively. In this paper, we will only be concerned with finite-dimensional quantum systems. In this case, we can choose a basis and think of $H(t)$ as a Hermitian matrix. We can split the Hamiltonian

$$H = H_d + \sum_{i=1}^{m} v_i(t)H_i,$$

where H_d is the part of Hamiltonian that is internal to the system and we call it the *drift* or *free Hamiltonian* and $\sum_{i=1}^{m} v_i(t)H_i$ is the part of Hamiltonian that can be externally changed. It is called the *control* or *rf Hamiltonian*. The equation for $U(t)$ dictates the evolution of the density matrix according to

$$\rho(t) = U(t)\rho(0)U^\dagger(t).$$

The problem we are ultimately interested in is to find the minimum time required to transfer the density matrix from the initial state ρ_0 to a final state ρ_F. Thus, we will be interested in computing the minimum time required to steer the system

$$\dot{U} = -i(H_d + \sum_{i=1}^{m} v_i H_i)\, U, \tag{1}$$

from identity, $U(0) = I$, to a final propagator U_F.

In the following section we establish a framework for studying such problems. For reasons suggested before our approach is more general than the current application requires, but this added generality does not complicate the development.

2   Preliminaries

We will assume that the reader is familiar with the basic facts about Lie groups and homogeneous spaces [1]. Throughout this paper, G will denote a compact semi-simple Lie group and e its identity

element (we use I to denote the identity matrix when working with the matrix representation of the group). As is well known there is a naturally defined bi-invariant metric on G, given by the Killing form. We denote this bi-invariant metric by $<,>_G$. Let K be a compact closed subgroup of G. We will denote by $L(G)$ the Lie algebra of right invariant vector fields on G and similarly $L(K)$ the Lie algebra of right invariant vector fields on K. There is a one to one correspondence between these vector fields and the tangent spaces $T_e(G)$ and $T_e(K)$, which we denote by $\mathfrak{g}$ and $\mathfrak{k}$ respectively. Consider the direct sum decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$ such that $\mathfrak{m} = \mathfrak{k}^{\perp}$ with respect to the metric.

To fix ideas, let $G = SU(n)$ and $\mathfrak{g} = \mathfrak{su}(n)$ be its associated Lie algebra of $n \times n$ traceless skew-Hermitian matrices. Then $< A, B >_G = tr(A^{\dagger}B)$, $A, B \in \mathfrak{su}(n)$ (which is proportional to the Killing metric) represents a bi-invariant metric on $SU(n)$.

It is well known that the (right) coset space $G/K = \{KU : U \in G\}$ (homogeneous space) admits the structure of a differentiable manifold [1]. Let $\pi : G \to G/K$ denote the natural projection map. Define $o \in G/K$ by $o = \pi(e)$. Given the decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, there exists a neighborhood of $0 \in \mathfrak{m}$ which is mapped homeomorphically onto a neighborhood of the origin $o \in G/K$ by the mapping $\pi \circ \exp|_{\mathfrak{m}}$. The tangent space plane $T_o(G/K)$ can be then identified with the vector subspace $\mathfrak{m}$. The geometry of homogeneous space will play an essential part in determining the shortest possible times for transfers.

The Lie group G acts on its Lie algebra $\mathfrak{g}$ by conjugation $Ad_G : \mathfrak{g} \to \mathfrak{g}$ (called the adjoint action). This is defined as follows. Given $U \in G$, $X \in \mathfrak{g}$, then

$$Ad_U(X) = \frac{d \ U^{-1} \exp(tX)U}{dt}\Big|_{t=0}.$$

Once again to fix ideas if $G = SU(n)$ and $U \in G$, $A \in \mathfrak{su}(n)$, then $Ad_U(A) = U^{\dagger}AU$. We use the notation

$$Ad_K(X) = \bigcup_{k \in K} Ad_k(X).$$

If the homogeneous space G/K is a Riemannian symmetric space [3], the Lie algebra decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$ (see [2] for properties of these orthogonal involutive Lie algebras) satisfies the commutation relation

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \ [\mathfrak{m}, \mathfrak{k}] \subset \mathfrak{m}, \ [\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{k}.$$

If $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$ contained in $\mathfrak{m}$, then $\mathfrak{h}$ is abelian because $[\mathfrak{m}, \mathfrak{m}] \in \mathfrak{g}$. It is well known [3] that:

Theorem 1 If $\mathfrak{h}$ and $\mathfrak{h}'$ are two maximal abelian subalgebras of $\mathfrak{m}$, then

1. There exists an element $\xi \in \mathfrak{h}$ whose centralizer in $\mathfrak{m}$ is just $\mathfrak{h}$.

2. There is an element $k \in K$ such that $Ad_k(\mathfrak{h}) = \mathfrak{h}'$.

3. $\mathfrak{m} = \bigcup_{k \in K} Ad_k(\mathfrak{h})$.

Thus the maximal abelian subalgebras of $\mathfrak{m}$ are all Ad_K conjugate and in particular they have the same dimension. The dimension will be called the *rank* of the symmetric space G/K and the maximal abelian subalgebras of $\mathfrak{m}$ are called the *Cartan subalgebras* of the pair $(\mathfrak{g}, \mathfrak{k})$. We will see in what follows that the structure of the time optimal control depends on the rank in an important way. We state a useful corollary of the above the theorem [3].

Corollary 1 Let G/K be a Riemannian symmetric space. Let $\mathfrak{h}$ be a Cartan subalgebra of the pair $(\mathfrak{g}, \mathfrak{k})$ and define $A = \exp(\mathfrak{h}) \subset G$. Then $G = KAK$.

Proof: $G = KP$, where $P = \exp(\mathfrak{m}) = \exp(\bigcup_{k \in K} Ad_k(\mathfrak{h})) = \bigcup_{k \in K} Ad_k(\exp(\mathfrak{h})) = \bigcup_{k \in K} Ad_k(A) \subset KAK$. Now $G = KKAK = KAK$. **Q.E.D**.

Note the space G/K is a union of maximal abelian subgroups $Ad_k(A)$, called *maximal tori.*

Assumption 1 Let $U \in G$ and let the control system

$$\dot{U} = [X_d + \sum_{i=1}^{m} v_i X_i] U, \; U(0) = e \tag{2}$$

be given. Please note we are working with the matrix representation of the group. We use $\{X_d, X_1, \ldots, X_m\}_{LA}$ to denote the Lie algebra generated by $\{X_d, X_1, \ldots, X_m\}$. We will assume that $\{X_d, X_1, \ldots, X_m\}_{LA} = \mathfrak{g}$, and since G is compact, it follows that the system (2) is controllable [4]. Let $\mathfrak{k} = \{X_i\}_{LA}$ and $K = \exp\{X_i\}_{LA}$ be the closed compact group generated by $\{X_i\}$. Given the direct sum decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$ where $\mathfrak{m} = \mathfrak{k}^\perp$ with respect to the bi-invariant metric $<,>_G$, let $X_d \in \mathfrak{m}$. We will assume that $Ad_K(\mathfrak{m}) \subset \mathfrak{m}$, in which case one says the homogeneous space G/K is reductive. All our examples will fall into this category.

Notation: Let $\mathcal{C}$ denote the class of all locally bounded measurable functions defined on the interval $[0, \infty)$ and taking value in $\mathbb{R}^m$. $\mathcal{C}[0, T]$ denotes their restriction on the interval $[0, T]$. We will assume throughout that in equation (2), $v = (v_1, v_2, \ldots, v_m) \in \mathcal{C}$. Given $v \in \mathcal{C}$, we use $U(t)$ to denote the solution of equation (2) such that $U(0) = e$. If, for some time $t \geq 0$, $U(t) = U'$, we say that the control v steers U into U' in t units of time and U' is attainable or reachable from U at time t.

Definition 1 (Reachable Set): The set of all $U' \in G$ attainable from U_0 at time t will be denoted by $R(U_0, t)$. Also we use the following notation

$$\mathbf{R}(U_0, T) = \bigcup_{0 \leq t \leq T} R(U_0, t)$$

$$\mathbf{R}(U_0) = \bigcup_{0 \leq t \leq \infty} R(U_0, t).$$

We will refer to $\mathbf{R}(U_0)$, as the reachable set of U_0.

Remark 1 From the right invariance of control systems it follows that $R(U_0, T) = R(e, T)U_0$, $\mathbf{R}(U_0, T) = \mathbf{R}(e, T)U_0$, and $\mathbf{R}(U_0) = \mathbf{R}(e)U_0$. Note that $\mathbf{R}(U_0, T)$ need not be a closed set, we use $\overline{\mathbf{R}(U_0, t)}$ to denote its closure.

Definition 2 (Infimizing Time): Given $U_F \in G$, we will define

$$t^*(U_F) = \inf \{t \geq 0 | \; U_F \in \overline{\mathbf{R}(e, t)}\}$$
$$t^*(KU_F) = \inf \{t \geq 0 | \; kU_F \in \overline{\mathbf{R}(e, t)}, \; k \in K\}$$

and $t^*(U)$ is called the *infimizing time.*

From a mathematical point of view, we may identify two goals in this paper: (1) to characterize $\overline{\mathbf{R}(e, t)}$ and hence compute $t^*(U_F)$, the infimizing time for $U_F \in G$, and (2) to characterize the infimizing control sequence v^n in (2), which in the limit $n \to \infty$, achieves the transfer time $t^*(U_F)$ of steering the system (2) from identity e to U_F. From the physics point of view, these results will help to establish the minimum time required and the optimal controls (the rf pulse sequence in NMR experiments) to achieve desired transfers in a spectroscopy experiment.
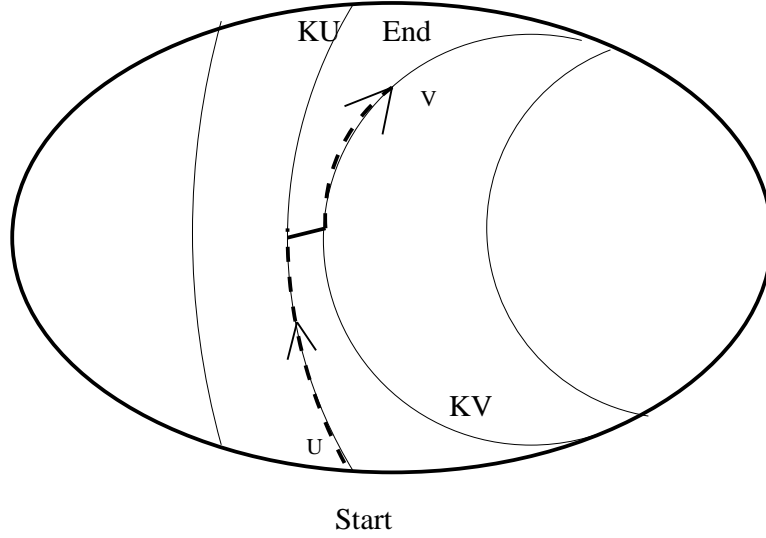
4

Figure 1: The panel shows the time optimal path between elements U and V belonging to G. The dashed line depicts the fast portion of the path corresponding to movement within the coset KU and, in traditional NMR language, corresponds to the pulse and the solid line corresponds to the slow portion of the curve connecting different cosets and corresponds to evolution of the couplings.

3   Time Optimal Control

The key observation is the following. In the control system (2), if $U_F \in K$ then $t^*(U_F) = 0$. To see this, note that by letting v in (2) be large, we can move on the subgroup K as fast as we wish. In the limit as v approaches infinity, we can come arbitrarily close to any point in K in arbitrarily small time with almost no effect from the term X_d. By same reasoning for any $U \in G$, $t^*(U) = t^*(kU)$ for $k \in K$. Thus, finding $t^*(U_F)$ reduces to finding the minimum time to steer the system (2) between the cosets Ke and KU_F.

This is illustrated in the Figure 1, where the cosets KU and KV are depicted and the infimizing time path between elements U and V belonging to G is shown. The dashed part of the curve illustrates the fast motion within the coset. The solid part of the curve corresponds to the drift part of the flow (also known as the evolution of couplings in NMR literature). The minimum time problem then corresponds to finding shortest path between these cosets or, in other words, the shortest path in the space G/K.

With this intuitive picture in mind, we now state some lemmas.

Lemma 1 Let $U \in G$ and $X : \mathbb{R} \to \mathfrak{g}$ be a locally bounded measurable function of time. If $X_n(t)$ converges to $X(t)$ in the sense that

$$\lim_{n \to \infty} \int_0^T \|X(t) - X_n(t)\| dt = 0,$$

then the solution of the differential equation $\dot{U} = X_n(t)U$ at time T converges to the solution of $\dot{U} = X(t)U$ at time T.

The proof of the above result is a direct consequence of the uniform convergence of the Peano-Baker series. We use this to show

Lemma 2 For the control system in equation (2), $t^*(U_F) = t^*(KU_F)$.

Proof: We first show that if $k \in K$, then $t^*(k) = 0$. Because $\exp\{X_1, \ldots, X_m\}_{LA} = K$, given any $T > 0$ there exists $\bar{v} \in \mathcal{C}(T)$, such that the solution of

$$\dot{U} = [\sum_{i=1}^{m} \bar{v}_i(t) X_i] U, \ U(0) = e$$

takes on the value k at time T. Now consider the family of control systems

$$\dot{U} = [X_d + \alpha \sum_{i=1}^{m} \bar{v}_i(\alpha t) X_i] U, \ U(0) = e.$$

Rescaling time as $\tau = \alpha t$, we obtain

$$\frac{dU}{d\tau} = [\frac{X_d}{\alpha} + \sum_{i=1}^{m} \bar{v}_i(\tau) X_i] U, \ U(0) = e.$$

Observe that, by Lemma 1, as $\alpha \to \infty$, $U(\tau)|_{\tau=T} = k$ or $\lim_{\alpha \to \infty} U(t)|_{t=\frac{T}{\alpha}} = k$. Therefore $k \in \overline{\mathbf{R}(e, T)}$, for all $T > 0$, implying $t^*(k) = 0$.

We now prove the general assertion. Let $t^*(U_F) = T$, we show that if $U_1 = kU_F$ for $k \in K$, then $t^*(U_1) = T$. Because $t^*(U_F) = T$, for any $T_1 > T$, $U_F \in \overline{R(e, T_1)}$, therefore there exists a family of control laws $v^r[0, T_1]$ such that the corresponding solutions $U^r(t)$ to the equation (2) satisfy $U^r(T_1) \to U_F$. From the first part of the proof, for any $T_2 > T_1$ there exists a control sequence $v^p[T_1, T_2]$ such that the solutions $U^p(t)$ to the family of control systems

$$\dot{U} = [X_d + \sum_{i=1}^{m} v_i^p X_i] U, \ U(T_1) = U_F$$

satisfies $U^p(t^p) \to U_1$, for $t^p < T_2$ and $t^p \to T_1$. Using the continuity of the solution of the differential equation to its initial condition and Lemma 1, we conclude that there exists a family of control laws $v^n[0, T_2]$ such that the corresponding solutions $U^n(t)$ to the family of control systems

$$\dot{U} = [X_d + \sum_{i=1}^{m} v_i^n X_i] U, \ U(0) = e$$

satisfy $U^n(t^n) \to U_1$, for $t^n < T_2$. Therefore, $U_1 \in \overline{R(e, T_2)}$. Since $T_2 > T_1$ is arbitrary $t^*(U_1) \leq T_1$. Because $T_1 > T$ is also arbitrary, we infer that $t^*(U_1) \leq T$. This shows that $t^*(U_1) \leq t^*(U_F)$. Now reverse the roles of U_F and U_1 to get the opposite inequality. This proves the claim. **Q.E.D**.

Remark 2 The above observation will help us make a bridge between the problem of computing $t^*(U_F)$ and the problem of computing minimum length paths for a related problem which we now explain.

Definition 3 (Adjoint Control System): Let $P \in G$. Associated with the control system (2) is the right invariant control system

$$\dot{P} = XP, \tag{3}$$

where now the control X no longer belongs to the vector space but is restricted to an adjoint orbit i.e., $X \in Ad_K(X_d) = \{k^{-1} X_d k | k \in K\}$. We call such a control system an *adjoint control system*.

For the control system (3), we say that $KU_F \in B(U_0, t')$ if there exists a control $X[0, t']$ which steers $P(0) = U_0$ to $P(t') \in KU_F$ in t' units of time. We use the notation

$$\mathbf{B}(U_0, T) = \bigcup_{0 \le t \le T} B(U_0, t).$$

From Lemma 1, we see that $\mathbf{B}(U_0, T)$ is closed.

We use

$$L^*(KU_F) = \inf \{t \ge 0 | \ KU_F \in \mathbf{B}(e, t)\}$$

to denote the minimum time required to steer the system (3) from identity e to the coset KU_F. We call it the *minimum coset time.*

Theorem 2 (Equivalence theorem): The infimizing time $t^*(U_F)$ for steering the system

$$\dot{U} = [X_d + \sum_{i=1}^{m} v_i X_i]U$$

from $U(0) = e$ to U_F is the same as the minimum coset time $L^*(KU_F)$, for steering the adjoint system

$$\dot{P} = XP, \ X \in Ad_K(X_d)$$

from $P(0) = e$ to KU_F.

Proof: Let $Q \in K$ satisfy the differential equation

$$\dot{Q} = [\sum_{i=1}^{m} v_i X_i]Q, \ Q(0) = e. \tag{4}$$

Let $P \in G$ evolve according to the equation

$$\dot{P} = (Q^{-1}X_d Q) \ P, \ P(0) = e. \tag{5}$$

Then observe that

$$\frac{d(Q \ P)}{dt} = [X_d + \sum_{i=1}^{m} v_i X_i](QP), \ \ Q(0)P(0) = e,$$

which is the same evolution equation as that of U, and since $U(0) = Q(0)P(0) = e$, by the uniqueness theorem for the differential equations, $U(t) = Q(t)P(t)$. Therefore, given a solution $\hat{U}(t)$ of equation (2) with the initial condition $\hat{U}(0)$, there exist unique curves $\hat{P}(t)$ and $\hat{Q}(t)$, defined through equations (4) and (5), satisfying $\hat{U}(t) = \hat{Q}(t)\hat{P}(t)$. Observe that if $\hat{U}(T) = U_F$ then it follows that $\hat{P}(T) \in KU_F$. If $U_F \in \overline{\mathbf{R}(e, T)}$, then there exists a sequence of control laws $v^r[0, T]$ such that the corresponding solutions $U^r(t)$ of (2) satisfy $U^r(T) \to U_F$. Therefore, the solutions $P^r(t)$ of the associated control system (4) satisfy $\lim_{r \to \infty} P^r(T) \in KU_F$. Because $\mathbf{B}(e, T)$ is closed, it follows that $KU_F \in \mathbf{B}(e, T)$, which implies that $L^*(KU_F) \le t^*(U_F)$.

To prove the equality observe that if $KU_F \in \mathbf{B}(e, T)$, then there exists a control $\bar{X}[0, T]$ such that the corresponding solution $\bar{P}(t)$ to (3) satisfies $\bar{P}(T) \in KU_F$. Because $\bar{X}(t) \in Ad_K(X_d)$, we can express $\bar{X}(t)$ as $\bar{Q}(t)^{-1}X_d\bar{Q}(t)$. It is well known [13] that we can find a family $v^r(t)$ of control laws such that the corresponding solution $Q^r(t)$ of

$$\dot{Q}^r = [\sum_{i=1}^{m} v_i^r X_i]Q^r, \ Q^r(0) = e$$

7

satisfy $\lim_{r \to \infty} \int_0^T \|\bar{Q}(t) - Q^r(t)\| dt = 0$. Hence, $\lim_{r \to \infty} \int_0^T \|\bar{X}(t) - (Q^r(t))^{-1} X_d Q^r(t)\| dt = 0$. Using Lemma 1, we claim that the solutions to family of differential equations

$$\dot{P}^r = [(Q^r)^{-1}(t) X_d Q^r(t)] P^r, \ P^r(0) = e$$

satisfies $\lim_{r \to \infty} P^r(T) \in KU_F$. Therefore, $t^*(KU_F) \leq T$. Since the choice of T was arbitrary, it follows $t^*(KU_F) \leq L^*(KU_F)$. Because $t^*(KU_F) = t^*(U_F)$, it follows that $t^*(U_F) \leq L^*(KU_F)$. Hence the proof. **Q.E.D**

Remark 3 The control system (2) evolves on the group G and induces a control system on the coset space G/K through the projection map π. The adjoint control system (3) is a representation of this induced control system. Observe that since $\|X\| = 1$ in (3), we can also define $L^*(KU_F)$ as the infimizing value of $\int_0^1 < \dot{P}, \dot{P} >^{\frac{1}{2}} dt$ for steering the system

$$\dot{P} = \gamma X P, \ \gamma > 0$$

from $P(0) = e$ to $P(1) \in KU_F$.

We will now compute $t^*(U_F)$ using the properties of the set $Ad_K(X_d)$. Based on the qualitative nature of time optimal trajectories of the system (2), we make the following classification.

1. **Riemannian Symmetric Case** In addition to Assumption 1, if we have the restriction $[\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{k}$, then we are in the Riemannian symmetric case as described in the section 2. We can further classify this case based on the rank of the symmetric space G/K.

   - **Pulse-drift-pulse sequence**(characteristic of single-spin systems) In this case, the rank of the symmetric space G/K is one. Roughly speaking the trajectories of the infimizing control sequence v^r (which in the limit $r \to \infty$, achieves the transfer time $t^*(U_F)$) converge to an impulse (which resembles an impulse of appropriate shape), followed by evolution under drift (for time $t^*(U_F)$) and a final impulse.

   - **Chained Pulse-drift-pulse sequence**(characteristic of two-spin system)In this case, the rank of the symmetric space G/K is more than one. The trajectories corresponding to an infimizing control sequence v^r in (3) converge to a chain of " impulse drift impulse" pattern. The infimizing time $t^*(U_F)$ is the time spent when the system just evolves under drift.

2. **Chatter sequence** In this case, G/K is no more a Riemannian symmetric case, i.e. $[\mathfrak{m}, \mathfrak{m}] \not\subset \mathfrak{k}$. This is a characteristic of more that two-spin systems.

   In this paper we will confine ourselves to the Riemannian symmetric case. The non-symmetric case will be treated in detail in a forthcoming paper.

Pulse-drift-pulse sequence

We begin with the first case where the rank of the symmetric space G/K is one. It follows from Theorem 1 that $\mathfrak{m} = \bigcup_{\alpha \geq 0} Ad_K(\alpha X_d)$. In this case, computing $t^*(U_F)$ reduces to finding the geodesic distance on the homogeneous space G/K. Given the bi-invariant metric $<,>_G$ on G, there is a corresponding left invariant metric $<,>_n$, on the homogeneous space G/K arising from the restriction of $<,>_G$ to $\mathfrak{m}$. Let $L_n(\gamma)$ represent the length of a curve $\gamma \in G/K$ under the standard

8

induced metric. There is a one-to-one correspondence between the curves $\{\gamma(t) \in G/K | \gamma(0) = 0, \ \gamma(1) = \pi(U_F), \ L_n(\gamma[0,1]) = T\}$ and the trajectories of system (3) satisfying $\{P(0) = e, P(T) \in KU_F\}$. Therefore, $L^*(KU)$ is the Riemannian distance between $\pi(e)$ and $\pi(U)$ under the standard metric $<,>_n$. This is computed in the following theorem, which characterizes geodesics on the homogeneous space G/K under the standard metric [1].

Theorem 3 Let G be a compact Lie group with a bi-invariant metric $<,>$, and K be a closed subgroup. Let $\mathfrak{g}$ and $\mathfrak{k}$ denote their Lie algebras with the direct sum decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, $\mathfrak{m} = \mathfrak{k}^\perp$. Consider the right invariant control system

$$\dot{U} = [X_d + \sum_{i=1}^{m} v_i X_i]U, \ U \in G, \ U(0) = e$$

where $v_i \in \mathbb{R}$, $X_d \in \mathfrak{m}$, and $\{X_i\}_{LA} = \mathfrak{k}$. Suppose G/K is a Riemannian symmetric space of rank one, then $t^*(U_F)$ is the smallest value of $\alpha > 0$ such that we can solve $U_F = Q_1 \exp(\alpha X_d) Q_2$ with $Q_1, Q_2 \in K$.

Proof: By the equivalence theorem $t^*(U_F) = L^*(KU_F)$, where $L^*(KU_F)$ is the minimum time for steering the system
$$\dot{P} = XP, \ X \in Ad_K(X_d)$$
from $P(0) = e$ to KU_F. Because G/K is a Riemannian symmetric space of rank one, $L^*(KU_F)$ is the Riemannian distance between o and $\pi(U)$ under the standard metric $<,>_n$. From [1], the geodesics under the metric $<,>_n$ originating from o take the form $\pi(\exp(\tau X))$ for $X \in \mathfrak{m}$. If $U_F = Q_1 \exp(tX_d) Q_2$ for $Q_1, \ Q_2 \in K$, then $\pi(U_F) = \pi(\exp(t \ Q_2^{-1} X_d Q_2))$. To see this note that $U_F = (Q_1 Q_2) Q_2^{-1} \exp(tX_d) Q_2 = (Q_1 Q_2) \exp(t \ Q_2^{-1} X_d Q_2)$. Thus, from the above assertion, $\pi(\exp(\tau \ Q_2^{-1} X_d Q_2))$ is the unique geodesic connecting o to $\pi(U_F)$ and has the length $L = t$. Hence the proof. **Q.E.D**

Remark 4 Roughly speaking, the time optimal trajectory (obtained as a limit of the infimizing sequence) for the system (2), which steers the system form $U(0) = e$ to $U_F = Q_1 \exp(\alpha X_d) Q_2$, takes the form $e \rightarrow Q_2 \rightarrow \exp(\alpha X_d) Q_2 \rightarrow Q_1 \exp(\alpha X_d) Q_2$, where the first and last step of this chain takes no time, and the time is required for the drift process(second step).

We now use illustrate these ideas through some examples.

Corollary 2 Let $U \in G = SU(2)$, and let $I_x = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $I_z = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ represent the Pauli spin matrices. Given the unitary evolution of the single-spin system

$$\dot{U} = -i[I_z + vI_x]U,$$

where the control $v \in \mathbb{R}$. Let $U_x = \exp(-iI_x t)$ represent the one-parameter subgroup generated by I_x. Given $U_F \in SU(2)$, there exists a unique $\beta \in [0, \ 2\pi]$ such that $U_F = U_1 \exp[-i\beta I_x]U_2$, where $U_1, U_2 \in U_x$. The infimizing time $t^*(U_F) = \beta$.

Proof: First note that the Lie algebra $\mathfrak{g} = \mathfrak{su}(2)$ has the decomposition $\mathfrak{m} = \{iI_y, iI_z\}$, $\mathfrak{k} = \{iI_x\}$, and $Ad_{U_x}(I_z) = \mathfrak{m}$. Observe from corollary 1 that any $U_F \in SU(2)$ has a representation $U_F =$

9

$Q_1 \exp[-i\alpha I_z]Q_2$, where $Q_1, Q_2 \in U_x$. Because $\exp[-itI_z]$ is periodic with period 4π, the β with the smallest absolute value for which $U_F = U_1 \exp[-i\beta I_z]U_2$ holds, belongs to the interval $[-2\pi, 2\pi]$. Because $-I_z \in Ad_{U_x}(I_z)$, we can restrict β to the interval $[0, 2\pi]$. The proof then follows directly from the Theorem 3. **Q.E.D.**

Corollary 3 Let $\Theta \in G = SO(3)$, and let $\Omega_x = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\Omega_z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$ represent the generators of rotation around $x-$ and $z-$axis. Consider the control system

$$\dot{\Theta} = [\Omega_z + v\Omega_x]\Theta,$$

where the control $v \in \mathbb{R}$. Let $\Theta_x = \exp(\Omega_x t)$ represent the one-parameter subgroup generated by Ω_x. Given $\Theta_F \in SO(3)$, there exists a unique $\beta \in [0, \pi]$ such that $\Theta_F = Q_1 \exp[\beta\Omega_x]Q_2$, where $Q_1, Q_2 \in \Theta_x$. The infimizing time $t^*(\Theta_F) = \beta$.

Proof: First note that the Lie algebra $\mathfrak{g} = \mathfrak{so}(3)$ has the decomposition $\mathfrak{m} = \{\Omega_y, \Omega_z\}$, $\mathfrak{k} = \{\Omega_x\}$, and $Ad_{\Theta_x}(\Omega_z) = \mathfrak{m}$. Observe that any $\Theta_f \in SO(3)$ has a representation $\Theta_f = Q_1 \exp[\alpha\Omega_z]Q_2$, where $Q_1, Q_2 \in \Theta_x$. Because $\exp[t\Omega_z]$ is periodic with period 2π, the proof is on the same lines as Corollary 2. **Q.E.D.**

We now generalize the example to the case where $G = SO(n)$, the group of $n \times n$ orthogonal matrices. The Lie algebra is $\mathfrak{g} = \mathfrak{so}(n)$, the set of $n \times n$ skew-symmetric matrices. The bi-invariant metric on G is $< \Omega, \Omega >= tr(\Omega^T\Omega)$. Consider the following decomposition of $\mathfrak{g}$. Let $\mathfrak{m}$ consists of skew-symmetric matrices which are zero except the first row and column and $\mathfrak{k}$ consists of skew symmetric matrices which are zero in the first row and column. Observe that $\mathfrak{k}$ generates the subgroup $SO(n-1)$. Then we have

Corollary 4 Let $\Theta \in G = SO(n)$ and let the control system

$$\dot{\Theta} = [\Omega_d + \sum_{i=1}^{m} v_i\Omega_i]\Theta, \ \Theta(0) = I$$

be given, where $\Omega_d \in \mathfrak{m}$, $v_i \in \mathbb{R}$, $\Omega_i \in \mathfrak{k}$ and $\Omega_d \in \mathfrak{m}$, such that $\exp[t\Omega_d]$ has period 2π. Suppose that $K = \exp\{\Omega_i\}_{LA} = SO(n-1)$. Given $\Theta_F \in SO(n)$, there exists a unique $\beta \in [0, \pi]$ such that $\Theta_F = \Theta_1 \exp[\beta\Omega_x]\Theta_2$, where $\Theta_1, \Theta_2 \in \Theta_x$. The infimizing time $t^*(\Theta_F) = \beta$.

Proof: Observe that $Ad_K(\Omega_d) = \mathfrak{m}$ and hence the proof is on the same lines as Corollary 2. **Q.E.D.**

Chained Pulse-drift-pulse sequence

Let us now consider the second case in our classification scheme. We now analyze the case when the rank of the Riemannian symmetric space G/K is greater than one.

Definition: (Weyl Orbit) Given the decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, let $\mathfrak{h} \subset \mathfrak{m}$ represent the maximal abelian subalgebra containing X_d. We use the notation $\Delta_{X_d} = \mathfrak{h} \bigcap Ad_K(X_d)$ to denote the maximal commuting set contained in the adjoint orbit of X_d. The set Δ_{X_d} is called the Weyl orbit of X_d.

We use $\mathfrak{c}(X_d) = \{\sum_{i=1}^{n} \beta_i X_i | \beta_i \geq 0, \sum \beta_i = 1, X_i \in \Delta_{X_d}\}$, to denote the convex hull of the Weyl orbit of X_d, with vertices given by the elements of the Weyl orbit of X_d.

We compute the infimizing time for the system (2), in the following Theorem (5), which is a generalization of the rank one case.

Remark 5 Recall from corollary (1) that, if $A = \exp(\mathfrak{h})$, where $\mathfrak{h}$ is the maximal abelian subalgebra contained in $\mathfrak{m}$, then $G = KAK$. Therefore given any $U_F \in G$, we can express $U_F = Q_1 \exp(Z)Q_2 = Q_1 Q_2 \exp(Ad_{Q_2}(Z))$, where Q_1, $Q_2 \in K$ and $Z \in \mathfrak{h}$. Suppose $Z = \sum_{i=1}^{n} \beta_i X_i$, $\beta_i \geq 0, X_i \in \Delta_{X_d}$. By choosing $X(t)$ to be $Ad_{Q_2}(X_i)$ for β_i units of time we can steer the adjoint control system $\dot{P} = X(t)P$ from the identity to the coset $KU_F = K \exp(Ad_{Q_2}(Z))$. The claim of the following theorem is that indeed the fastest way to get to the coset KU_F is to flow on the maximal torus, $Ad_{Q_2}(A)$, $Q_2 \in K$, containing the coset KU_F.

We now state a convexity theorem due to Kostant [19], which is the main idea behind the following theorem.

Theorem 4 (Kostant's Convexity Theorem): Given the direct sum decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, let $\mathfrak{h} \subset \mathfrak{m}$ represent a maximal abelian subalgebra containing $X_d \in \mathfrak{m}$. Let $\Gamma : \mathfrak{m} \rightarrow \mathfrak{h}$, be the orthogonal projection of $\mathfrak{m}$ onto $\mathfrak{h}$. Then $\Gamma : Ad_K(X_d) = \mathfrak{c}(X_d)$, where $\mathfrak{c}(X_d)$ is the convex hull of the Weyl orbit of X_d as defined above.

Theorem 5 (Time Optimal Tori Theorem:) Let G be a compact matrix Lie group and K be a closed subgroup with $\mathfrak{g}$ and $\mathfrak{k}$ their Lie algebras, respectively such that G/K is a Riemannian symmetric space. Let the direct sum decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, such that $\mathfrak{m} = \mathfrak{k}^{\perp}$, be given. Consider the right invariant control system

$$\dot{U} = [X_d + \sum_{i=1}^{m} v_i X_i]U, \ U \in G, \ U(0) = e,$$

where $v_i \in \mathbb{R}$, $X_d \in \mathfrak{m}$, $\{X_i\}_{LA} = \mathfrak{k}$. Then any $U_F = Q_1 \exp(\alpha Y)Q_2$, where $\alpha > 0, Q_1$, $Q_2 \in K$, and $Y \in \mathfrak{c}(X_d)$, belongs to the closure of the reachable set. The infimizing time $t^*(U_F)$ is the smallest value of $\alpha > 0$, such that we can solve

$$U_F = Q_1 \exp(\alpha Y)Q_2,$$

where Q_1, $Q_2 \in K$ and Y belongs to the convex hull $\mathfrak{c}(X_d)$.

We sketch here the outline of a proof. A rigorous proof from a control theoretic viewpoint will be presented elsewhere.

Once again, we will compute $t^*(U_F)$ by finding the minimum time it takes to steer the system $\dot{P} = XP$ from $X(0) = e$ to the coset KU_F. Following the notation of the corollary 1, let $A = \exp(\mathfrak{h})$ denote the maximal torus contained inside G/K. From corollary 1, we know that $G/K = Ad_K(A)$. Thus it suffices to prove the theorem for $U_F \in A$, therefore we will show that if $U_F \in A$ then T is the smallest value of $\sum_{i=1}^{m} \alpha_i$, $\alpha_i \geq 0$ such that $U_F = \exp(\sum_{i=1}^{m} \alpha_i X_i)$, where $X_i \in \Delta_{X_d}$, then $t^*(U_F) = T$. It is immediate that in the adjoint control system $\dot{P} = XP$, by letting X to be X_i for α_i units of time we can reach U_F in T units of time. All we need to show is that we can reach the coset KU_F no sooner. Let $\bar{P}(t)$ be the shortest (or time optimal) trajectory of the adjoint control system $\dot{P} = XP$ that steers $P(0) = e$ to the coset KU_F. Let $\pi_A : G/K \rightarrow A$, denote the projection

11

on the maximal subgroup, whereby for $A_1 \in A$, $\pi_A : k^{-1}A_1 k \to A_1$ (note that the projection is only unique modulo a Weyl group action, hence an explicit choice is required to make the projection unique). Let $a(t) \in A$ be a continuous path obtained from the projection of $\bar{P}(t)$, onto A. Then observe the projection π_A induces the map $\pi_{A_*} : \dot{\bar{P}}(t) \to \dot{a}(t)$. Then the evolution of the curve $a(t)$ can be expressed as $\dot{a}(t) = \Omega\, a(t)$, where $\Omega = \Gamma(Ad_{\tilde{k}}(X_d))$, for some $\tilde{k} \in K$ (recall $\Gamma : \mathfrak{m} \to \mathfrak{h}$ is the orthogonal projection onto $\mathfrak{h}$). Now using Kostant's convexity theorem, we have $\Omega \in \mathfrak{c}(X_d)$. Therefore we can write $\dot{a}(t) = (\sum_{i=1}^m \beta_i X_i)\, a(t)$, where $X_i \in \Delta_{X_d}$ and $\sum_{i=1}^m \beta_i = 1$ for $\beta_i \geq 0$. This implies that if $T = \sum_{i=1}^m \alpha_i$ is the smallest value for which $U_F \in A$ satisfies $U_F = \exp(\sum_{i=1}^m \alpha_i X_i)$ for $\alpha_i \geq 0$, then the path $a(t)$ will atleast takes T units of time to reach U_F.

Remark 6 The theorem characterizes $\mathbf{B}(e, t)$, the reachable set for the adjoint system. This is given by
$$K\mathbf{B}(e, t) = K \exp(\alpha Y)K, \ 0 \leq \alpha \leq t$$
where Y belongs to the convex hull $\mathfrak{c}(X_d)$.

4   Spin Algebra

The Lie Group which we will be most interested in is $SU(2^n)$, the special unitary group describing the evolution of n coupled spins $\frac{1}{2}$. Its Lie algebra $\mathfrak{su}(2^n)$ is a $4^n - 1$ dimensional space of traceless $n \times n$ skew-Hermitian matrices. The orthonormal basis which we will use for this space is expressed as tensor products of Pauli spin matrices [12]

$$
\begin{aligned}
I_x &= \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
I_y &= \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
I_z &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\end{aligned}
$$

The matrices (I_x, I_y, I_z) are the generators for rotation in the two dimensional Hilbert space and basis for the Lie algebra of traceless skew-Hermitian matrices $\mathfrak{su}(2)$. They obey the well known commutation relations
$$[I_x\ I_y] = iI_z \ ; \ [I_y\ I_z] = iI_x \ ; \ [I_z\ I_x] = iI_y.$$

Then the basis for $\mathfrak{su}(2^n)$ takes the form $\{iB_s\}$ where

$$B_s = 2^{q-1} \prod_{k=1}^n (I_{k\alpha})^{a_{ks}}, \tag{6}$$

$\alpha = x, y, or\ z$ and

$$I_{k\alpha} = \mathbf{1}\ \otimes \cdots \otimes I_\alpha \otimes \mathbf{1}\ , \tag{7}$$

where I_α the Pauli matrix appears in the above expression only at the kth position, and $\mathbf{1}$ the two dimensional identity matrix, appears everywhere except at the kth position. a_{ks} is 1 for q of the

indices and 0 for the remaining. Note that $q \geq 1$ as $q = 0$ corresponds to the identity matrix and is not a part of the algebra. As an example for $n = 2$ the basis for $\mathfrak{su}(4)$ takes the form

$$
\begin{array}{ll}
q = 1 & I_{1x}, I_{1y}, I_{1z}, I_{2x}, I_{2y}, I_{2z} \\
q = 2 & 2I_{1x}I_{2x}, 2I_{1x}I_{2y}, 2I_{1x}I_{2z} \\
& 2I_{1y}I_{2x}, 2I_{1y}I_{2y}, 2I_{1y}I_{2z} \\
& 2I_{1z}I_{2x}, 2I_{1z}I_{2y}, 2I_{1z}I_{2z}.
\end{array}
$$

It is important to note that these operators are only normalized for $n = 2$ as

$$
tr(B_r B_s) = \delta_{rs} 2^{n-2}.
$$

To fix ideas, lets compute one of these operators explicitly for $n = 2$

$$
I_{1z} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
$$

which takes the form

$$
I_{1z} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.
$$

We will often refer to the algebra of $\mathfrak{su}(2^n)$ as the *spin algebra*.

5   Optimal Transfer in Two-Spin Systems

In this section, we will apply our general results on the time optimal control for the specific case of a heteronuclear two-spin system. In particular, we consider the following important heteronuclear two-spin system discussed in detail in [6]. By going to a rotating frame, the free evolution part of the Hamiltonian has been reduced to just a scalar coupling evolution. The system then takes the following form.

Let $U \in SU(4)$, which evolves as

$$
\dot{U} = -i(H_d + \sum_{i=1}^{4} u_i H_i)U, \tag{8}
$$

where

$$
\begin{array}{rcl}
H_d & = & 2\pi J I_z S_z \\
H_1 & = & 2\pi I_x \\
H_2 & = & 2\pi I_y \\
H_3 & = & 2\pi S_x \\
H_4 & = & 2\pi S_y,
\end{array}
$$

where I_x, I_y and I_z represent operators for the first spin and have the same meaning as I_{1x}, I_{1y} and I_{1z}, respectively, as explained in previous section 4. Similarly S_x, S_y, and S_z represent operators for the second spin and have the same meaning as I_{2x}, I_{2y} and I_{2z}. The symbol J represents the strength of the scalar coupling between I and S. Observe that the subgroup K generated by $\{H_i\}_{i=1}^4$ is $SU(2) \times SU(2)$.

We first compute the infimizing time for steering the system (8).

Theorem 6 For the heteronuclear spin system, described by the equation (8), the infimizing time $t^*(U_F)$ is the smallest value of $\sum_{i=1}^3 \alpha_i$, $\alpha_i > 0$, such that we can solve

$$U_F = Q_1 \exp(-i2\pi J(\alpha_1 I_x S_x + \alpha_2 I_y S_y + \alpha_3 I_z S_z))Q_2,$$

where Q_1 and Q_2 belong to K.

Proof: Consider the direct sum decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$, where $\mathfrak{m} = span\{I_\alpha S_\beta\}$, $\mathfrak{k} = span\{I_\alpha, S_\beta\}$, and $(\alpha, \beta) \in (x, y, z)$. Then observe $[\mathfrak{m}, \mathfrak{m}] \in \mathfrak{k}$, $[\mathfrak{m}, \mathfrak{k}] \in \mathfrak{m}$, and $[\mathfrak{k}, \mathfrak{k}] \in \mathfrak{k}$. Furthermore, observe that $\Delta_{I_z S_z} = \{\pm I_z S_z, \pm I_x S_x, \pm I_y S_y\}$, therefore for $\alpha \geq 0$, we have $Ad_K(\alpha \, \mathfrak{c}(X_d))) = \mathfrak{m}$. Thus the above example satisfies all the conditions of the theorem 5. Hence the proof. **Q.E.D**

Now we address the question of maximum possible achievable transfer in some given time T. For this purpose we define the transfer efficiency.

Definition 4 (Transfer Efficiency): Given the evolution of the density matrix $\rho(t) = U(t)\rho(0)U^\dagger(t)$, where

$$\dot{U} = -i(\ H_d + \sum_{i=1}^m u_i H_i\ )U,\ U(0) = I,$$

define the *transfer efficiency* $\eta(t)$ to some given target operator F as

$$\eta(t) = \|tr(F^\dagger U(t)\rho(0)U^\dagger(t))\|.$$

Remark 7 *In the formula for the transfer efficiency, we always assume that the starting operator $\rho(0)$ and the final operator F are both normalized to have norm one (i.e. $tr(F^\dagger F) = 1$).*

We will now look at the in-phase and anti-phase transfers in the two-spin system, whose evolution is given by equation (8). We give here expressions for maximum transfer efficiencies. We first prove some lemmas, which will be required in computing transfer efficiencies.

Lemma 3 let $p = \begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix}$ and let Σ be a real diagonal matrix

$$\Sigma = \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix}.$$

If $a_i \geq a_j \geq a_k \geq 0$, where $\{i, j, k\} \in \{1, 2, 3\}$ and let $U, V \in SO(3)$, then the maximum value of $\|p^\dagger U \Sigma V p\|$ is $a_i + a_j$.

14

Proof: Let

$$\Lambda = \begin{bmatrix} \sqrt{a_1} & 0 & 0 \\ 0 & \sqrt{a_2} & 0 \\ 0 & 0 & \sqrt{a_3} \end{bmatrix}.$$

By definition $\Sigma = \Lambda^\dagger \Lambda$. Using Cauchy Schwartz inequality $\|p^\dagger U \Sigma V p\| \leq \|\Lambda V p\| \, \|\Lambda U p\|$. Observe, the maximum value of $\|\Lambda V p\|$ is $\sqrt{a_i + a_j}$. Therefore $\|p^\dagger U \Sigma V p\| \leq a_i + a_j$. Clearly for appropriate choice of U and V, this upper bound is achieved (For example, in case $a_1 \geq a_2 \geq a_3$, the bound is achieved for U and V identity). Hence the result follows. **Q.E.D**.

Lemma 4 Consider the function $f(\alpha_1, \alpha_2, \alpha_3) = \sin(J\pi\alpha_1)\sin(J\pi\alpha_2) + \sin(J\pi\alpha_1)\sin(J\pi\alpha_3)$. If $\alpha_1, \alpha_2, \alpha_3 \geq 0$ and $\alpha_1 + \alpha_2 + \alpha_3 = T$, where $T \leq \frac{3}{2J}$, then the maximum value of $f(\alpha_1, \alpha_2, \alpha_3)$ is $2sin(J\pi a)sin(J\pi b)$, where $a + 2b = t$ and $\tan(J\pi a) = 2\tan(J\pi b)$.

Proof: Let

$$H(\alpha_1, \alpha_2, \alpha_3, \lambda) = \sin(J\pi\alpha_1)\sin(J\pi\alpha_2) + \sin(J\pi\alpha_1)\sin(J\pi\alpha_3) + \lambda(\alpha_1 + \alpha_2 + \alpha_3 - T).$$

The necessary condition for optimality gives $\frac{\partial H}{\partial \alpha_1} = 0$, $\frac{\partial H}{\partial \alpha_2} = 0$, $\frac{\partial H}{\partial \alpha_3} = 0$, which imply respectively that

$$\pi J(\cos(J\pi\alpha_1)\sin(J\pi\alpha_2) \quad + \quad \cos(J\pi\alpha_1)\sin(J\pi\alpha_3)) + \lambda = 0 \tag{9}$$
$$\pi J(\sin(J\pi\alpha_1)\cos(J\pi\alpha_2)) \quad + \quad \lambda = 0 \tag{10}$$
$$\pi J(\sin(J\pi\alpha_1)\cos(J\pi\alpha_3)) \quad + \quad \lambda = 0 \tag{11}$$

From equation (10) and (11), we obtain that either $\sin(J\pi\alpha_1) = 0$ or $\cos(J\pi\alpha_2)) = \cos(J\pi\alpha_3))$. The first condition does not give a maxima as it makes f identically zero. The second condition implies

$$J\pi\alpha_2 = 2m\pi + J\pi\alpha_3. \tag{12}$$

Since $\alpha_2, \alpha_3 \geq 0$ and $\alpha_2 + \alpha_3 \leq T \leq \frac{3}{2J}$, condition (12) is only satisfied for $m = 0$. Therefore, $\alpha_1 = \alpha_2$. Now substituting this in (9) and using the equations (9) and (10), we get the desired result **Q.E.D**.

Theorem 7 (Maximum in-phase transfer) Consider the evolution for the heteronuclear IS spin system as defined by Equation (8). Let $\rho(0) = \frac{S_x - iS_y}{\sqrt{2}}$ and $F = \frac{I_x - iI_y}{\sqrt{2}}$. For $t \leq \frac{3}{2J}$, the maximum achievable transfer

$$\eta^*(t) = sin(J\pi a)sin(J\pi b),$$

where $a + 2b = t$ and $\tan(J\pi a) = 2\tan(J\pi b)$. For $t \geq \frac{3}{2J}$ the maximum achievable transfer is one.

Proof: Let

$$\Lambda(\alpha_1, \alpha_2, \alpha_3) = \exp(-i2\pi J(\alpha_1 I_x S_x + \alpha_2 I_y S_y + \alpha_3 I_z S_z)).$$

From now on we will simply write $\Lambda(\alpha_1, \alpha_2, \alpha_3)$ as Λ. From Theorem 6, any unitary propagator U_F belonging to the set

$$\mathbf{R}(e, t) = \{Q_1 \Lambda Q_2 | \ Q_1, \ Q_2 \in K \ \ \alpha_i > 0, \ \sum_{i=1}^{3} \alpha_i \leq t\},$$

can be produced by appropriate pulse sequence in (8). Therefore we will maximize $\|tr(F^\dagger U(t)\rho(0)U^\dagger(t))\|$, for $U(t) \in \mathbf{R}(e, t)$. Let $I = \exp\{iI_x, iI_y, iI_z\}$ and $S = \exp\{iS_x, iS_y, iS_z\}$. By definition, $K = S \times I$. In the expression

$$\eta(t) = \|tr(Q_1^\dagger F^\dagger Q_1 \Lambda Q_2 \rho(0) Q_2^\dagger \Lambda^\dagger)\|,$$

$\rho(0)$ commutes with I and F commutes with S, therefore it suffices to restrict Q_1 and Q_2 to I and S, respectively.

Let $\mathfrak{s}$ denote the subspace spanned by the orthonormal basis $\{S_x, S_y, S_z\}$ and $\mathfrak{i}$ denote the subspace spanned by the orthonormal basis $\{I_x, I_y, I_z\}$. We represent the starting operator $\rho(0) = \frac{1}{\sqrt{2}}(S_x - iS_y)$ as a column vector $p = \frac{1}{\sqrt{2}}[1 \ -i \ 0]^T$ in $\mathfrak{s}$. The action $\rho(0) \to Q_2\rho(0)Q_2^\dagger$ can then be represented as $p \to Vp$ where V is a orthogonal matrix.

Let P_I denote the projection on the subspace $\mathfrak{i}$. A simple computation yields that

$$
\begin{aligned}
P_I(\Lambda S_x \Lambda^\dagger) &= \sin(J\pi\alpha_2)\sin(J\pi\alpha_3)I_x \\
P_I(\Lambda S_y \Lambda^\dagger) &= \sin(J\pi\alpha_1)\sin(J\pi\alpha_3)I_y \\
P_I(\Lambda S_z \Lambda^\dagger) &= \sin(J\pi\alpha_2)\sin(J\pi\alpha_3)I_z.
\end{aligned}
$$

We denote the target operator $F = \frac{1}{\sqrt{2}}(I_x - iI_y)$ as a column vector $\frac{1}{\sqrt{2}}[1 \ -i \ 0]^T$ in $\mathfrak{i}$. The action $\rho(0) \to P_I(\Lambda Q_2\rho(0)Q_2^\dagger\Lambda^\dagger)$ can be written as $p \to \Sigma Vp$, where

$$
\Sigma = \begin{bmatrix}
\sin(J\pi\alpha_2)\sin(J\pi\alpha_3) & 0 & 0 \\
0 & \sin(J\pi\alpha_1)\sin(J\pi\alpha_3) & 0 \\
0 & 0 & \sin(J\pi\alpha_1)\sin(J\pi\alpha_2)
\end{bmatrix},
$$

Therefore we can rewrite $\eta(t) = \|tr(Q_1^\dagger F^\dagger Q_1 \Lambda Q_2\rho(0)Q_2^\dagger\Lambda^\dagger)\|$ as $\eta(t) = \|p^\dagger U\Sigma Vp\|$, where U and V are real orthogonal matrices. Using the result of Lemma (3), we get that for $\sin(J\pi\alpha_1) \geq \sin(J\pi\alpha_2) \geq \sin(J\pi\alpha_3) \geq 0$, the maximum value of $\eta(t)$ is

$$
\frac{\sin(J\pi\alpha_1)\sin(J\pi\alpha_2) + \sin(J\pi\alpha_1)\sin(J\pi\alpha_3)}{2}.
$$

Now we maximize the above expression with respect to $\alpha_1, \alpha_2, \alpha_3$ as worked out in Lemma 4 to get the above result.

Now we prove the last part of the theorem. Note for $t = \frac{3}{2J}$, the maximum achievable transfer is one. Because $\rho(0)$ and F are normalized, this is the maximum possible transfer between these operators. If $t > \frac{3}{2J}$, say $t = T + \frac{3}{2J}$, we can always arrange matters so that $U(T) = e$ (by creating a propagator $U(T/2) = exp(-i2\pi J(\frac{T}{2}I_zS_z))$ and then creating its inverse $exp(i2\pi J(\frac{T}{2}I_zS_z))$ from $T/2$ to T). In the remaining $\frac{3}{2J}$ units of time, we can produce the optimal propagator.

Q.E.D

The optimal transfer curve is plotted in comparison with the transfer achieved using the isotropic mixing Hamiltonian in the Figure 2. The unitary propagator $U(t)$ in the isotropic mixing Hamiltonian case takes the form

$$
U(t) = \exp(-i\frac{2\pi Jt}{3}(I_zS_z + I_xS_x + I_yS_y)).
$$

For small mixing times the transfer amplitude achieved by the optimal experiment is up to 12.5 % larger than the transfer achieved by isotropic mixing. This is a previously unknown result that will find immediate practical applications in NMR spectroscopy.

Theorem 8 (Maximum anti-phase transfer) Consider the evolution for the heteronuclear IS spin system as defined by equation (8). Let $\rho(0) = \sqrt{2}I_zS^- = \sqrt{2}I_z(S_x - iS_y)$ and $F = I^- = \frac{I_x - iI_y}{\sqrt{2}}$.
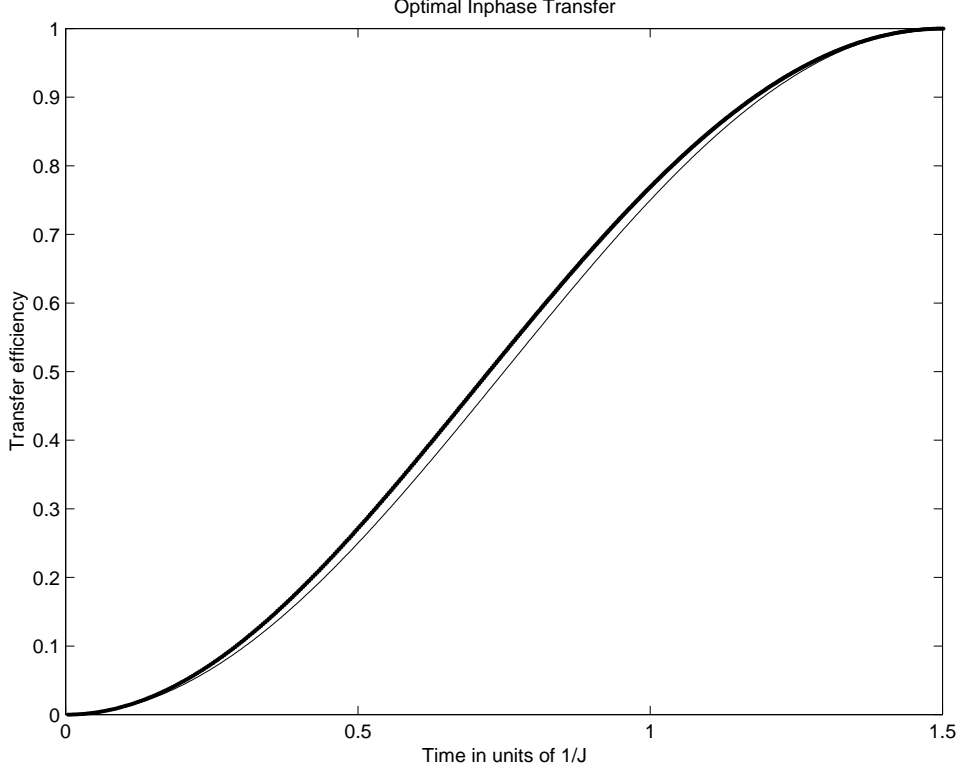
16

Figure 2: The panel shows the comparison between the best achievable transfer (bold curve) and the transfer achieved using the isotropic mixing Hamiltonian for the in-phase transfer in 2 spin case. On X axis is plotted time in units of $1/J$.

Then, for $t \leq 1/J$, the maximum achievable transfer $\eta^*(t)$ is

$$\|tr(F^\dagger U(t)\rho(0)U^\dagger(t))\| = sin(J\pi t/2).$$

For $t \geq \frac{1}{J}$, the maximum achievable transfer is one.

Proof: Let

$$\Lambda = \exp(-i2\pi J(\alpha_1 I_x S_x + \alpha_2 I_y S_y + \alpha_3 I_z S_z)).$$

From theorem 6

$$U(t) \in \{Q_1 \Lambda Q_2 | \ Q_1, \ Q_2 \in K \ \alpha_i > 0, \ \sum_{i=1}^{3} \alpha_i \leq t\}.$$

Let $S = \exp\{iS_x, iS_y, iS_z\}$ and $I = \exp\{iI_x, iI_y, iI_z\}$. By definition $K = S \times I$. In the expression for

$$\eta = \|tr(F^\dagger Q_1 \Lambda Q_2 \rho(0) Q_2^\dagger \Lambda Q_1^\dagger)\|,$$

let $Q_2 = Q_{2I} \times Q_{2S}$, where $Q_{2I} \in I$ and $Q_{2S} \in S$. Let the optimal $Q_2^* = Q_{2I}^* \times Q_{2S}^*$ be such that

$$Q_{2I}^* \rho(0) Q_{2I}^{*\dagger} = Q_{2I}^* I_z S^- Q_{2I}^{*\dagger} = a_z I_z S^- + a_y I_y S^- + a_x I_x S^-,$$

where $a_x^2 + a_y^2 + a_z^2 = 1$. Denote

$$
\begin{aligned}
\eta_z &= \|tr(F^\dagger (Q_1 \Lambda Q_{2S})^*(\sqrt{2}I_z S^-)(Q_{2S}^\dagger \Lambda^\dagger Q_1^\dagger)^*)\| \\
\eta_y &= \|tr(F^\dagger (Q_1 \Lambda Q_{2S})^*(\sqrt{2}I_y S^-)(Q_{2S}^\dagger \Lambda^\dagger Q_1^\dagger)^*)\| \\
\eta_x &= \|tr(F^\dagger (Q_1 \Lambda Q_{2S})^*(\sqrt{2}I_x S^-)(Q_{2S}^\dagger \Lambda^\dagger Q_1^\dagger)^*)\|.
\end{aligned}
$$

17

Then observe that

$$\eta(t) \leq a_z \eta_z + a_y \eta_y + a_x \eta_x.$$

We first compute the maximum of η_z. Let P_I denote the projection on the subspace generated by $\{I_x, I_y, I_z\}$, then a simple computation yields

$$
\begin{aligned}
P_I(\Lambda I_z S_x \Lambda^\dagger) &= \frac{1}{2}\sin(J\pi\alpha_1)I_y \\
P_I(\Lambda I_z S_y \Lambda^\dagger) &= \frac{1}{2}\sin(J\pi\alpha_2)I_x \\
P_I(\Lambda I_z S_z \Lambda^\dagger) &= 0.
\end{aligned}
$$

Since $\{I_z S_x, I_z S_y, I_z S_z\}$ forms an orthogonal pair, we can rewrite

$$\|tr(F^\dagger Q_1 \Lambda Q_{2S}(\sqrt{2}I_z S^-)(Q_{2S}^\dagger \Lambda^\dagger Q_1^\dagger))\|$$

as

$$\eta(t) = \|p^\dagger U \Sigma V p\|,$$

where $p = [1 \ -i \ 0]^T$,

$$
\Sigma = \begin{bmatrix}
\frac{\sin(J\pi\alpha_2)}{2} & 0 & 0 \\
0 & \frac{\sin(J\pi\alpha_1)}{2} & 0 \\
0 & 0 & 0
\end{bmatrix},
$$

and U and V are real orthogonal matrices. From Lemma 3, it follows that the maximum value of η_z is

$$\frac{\sin(J\pi\alpha_2)}{2} + \frac{\sin(J\pi\alpha_1)}{2}.$$

We can compute the maximum of the above expression under the constraint $\alpha_1 + \alpha_2 = t \leq 1/(J)$. The maximum value of the above expression is obtained for $\alpha_1 = \alpha_2$. The maximum value is $\sin(J\pi t/2)$ for $t \leq 1/J$. Similarly, the maximum value of η_x and η_y is as above. Since $a_x^2 + a_y^2 + a_z^2 = 1$, we get the desired result.

The final proposition of the theorem has the same proof as in Theorem 6   **Q.E.D**.

The optimal transfer curve for the anti-phase transfer plotted as a function of time measured in units of $1/J$ is shown in the Figure 3.This clearly shows that the transfer efficiency achieved using the known mixing sequence [10] is optimal.

6   Conclusion

In this paper, we presented a mathematical formulation of the problem of finding shortest pulse sequences in coherent spectroscopy. We showed how the problem of computing minimum time to produce a unitary propagator can be reduced to finding shortest length paths on certain coset spaces. A remarkable feature of time optimal control laws is that they are singular, i.e. the control is zero most of the time, with impulses in-between. We explicitly computed the shortest transfer times and maximum achievable transfer in a given time for the case of heteronuclear two-spin transfers. In a forthcoming paper, we plan to extend these results to higher spin systems.

References

[1] S. Kobayashi, and K. Nomizu. *Foundations of Differential Geometry Vol.1 and 2* (Interscience Publishers) (1969).
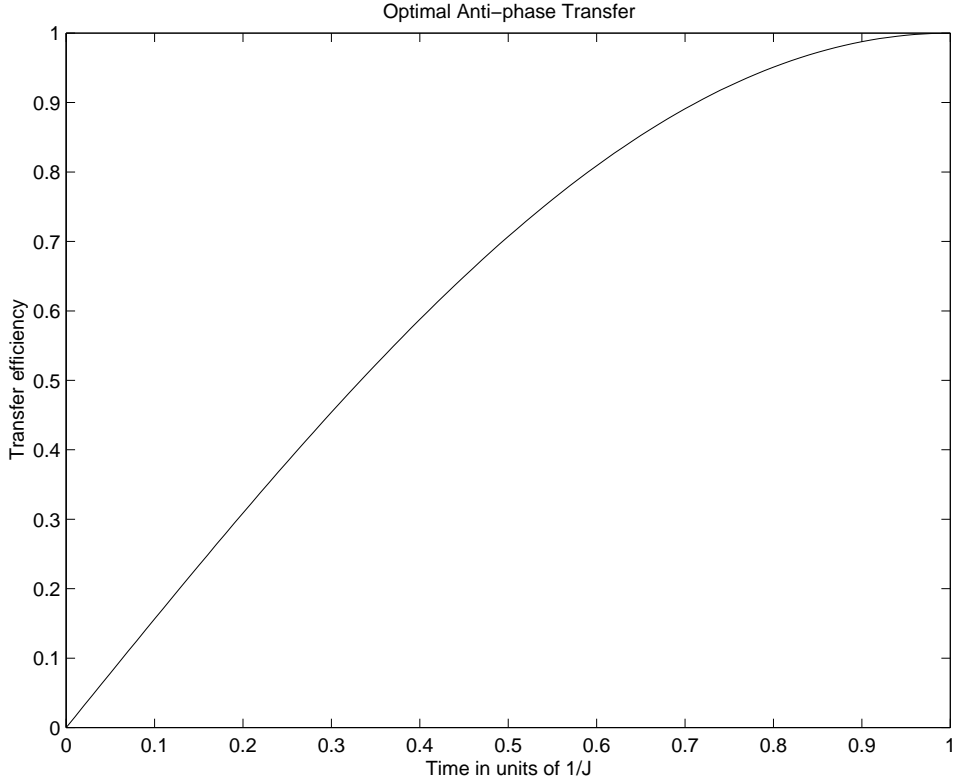
Figure 3: The panel shows the best achievable transfer as a function of time measured in units of $1/J$ for the anti-phase transfer in 2 spin case.

[2] S. Helgason *Differential Geometry, Lie Groups, and Symmetric Spaces* (Academic Press) (1978).

[3] J A. Wolf, *Spaces of Constant Curvature* (Publish or Perish, Inc., 1984).

[4] V. Jurdjevic and H. J. Sussmann *Journal of Differential Equations* **12**, 313 (1972).

[5] R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Oxford University Press, Oxford, 1987).

[6] S. J. Glaser, T. Schulte-Herbrüggen, M. Sieveking, O. Schedletzky, N. C. Nielsen, O. W. Sørensen and C. Griesinger, *Science* **280**, 421 (1998).

[7] W. S. Warren, H. Rabitz, M. Dahleh, *Science* **259**, 1581 (1993).

[8] N. A. Gershenfeld and I. L. Chuang, *Science* **275**, 350 (1997);

D. G. Cory, A. Fahmy, T. Havel, *Proc. Natl. Acad. Sci. USA* **94**, 1634 (1997).

[9] T. Untidt, T. Schulte-Herbrüggen, B. Luy, S. J. Glaser, C. Griesinger, O. W. Sørensen and N. C. Nielsen, *Molecular Physics* **95**, 787 (1998);

T. Untidt, S. J. Glaser, C. Griesinger and N. C. Nielsen, *Molecular Physics* **96**, 1739 (1999).

[10] J. Cavanagh, A. G. Palmer III, P. E. Wright, M. Rance, *J. Magn. Reson.* **91**, 429 (1991);

A. G. Palmer III, J. Cavanagh, P. E. Wright, M. Rance, *ibid.* **93**, 151 (1991);

L. E. Kay; *J. Am. Chem. Soc.* **115**, 2055 (1993);

M. Sattler, P. Schmidt, J. Schleucher, O. Schedletzky, S. J. Glaser and C. Griesinger, *J. Magn. Reson.* **B 108**, 235 (1995);

J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky, S. J. Glaser, O. W. Sørensen and C. Griesinger, *J. Biomol. NMR* **4**, 30 (1994).

[11] D. P. Weitekamp, J. R. Garbow, A. Pines, *J. Chem. Phys.* **77**, 2870 (1982), ibid. **80**, 1372 (1984);

P. Caravatti, L. Braunschweiler, R. R. Ernst, *Chem. Phys. Lett.* **100**, 305 (1983);

S. J. Glaser and J. J. Quant, in *Advances in Magnetic and Optical Resonance*, W. S. Warren, Ed. (Academic Press, New York, 1996), vol. 19, pp. 59-252.

[12] O.W. Sørensen, G.W. Eich, M.H. Levitt, G.Bodenhausen and R.R. Ernst, *Progr. NMR Spectrosc.* **16**, 163 (1983).

[13] G.W. Haynes and H. Hermes, *SIAM J. Control and Optimization* **8**, 450 (1970).

[14] R.W. Brockett and Navin Khaneja, "Stochastic Control of Quantum Ensembles" *System Theory: Modeling, Analysis and Control* (Kluwer Academic Publishers, Inc., 1999).

[15] Navin Khaneja, *Geometric Control in Classical and Quantum Systems* (Ph.d. Thesis, Harvard University, 2000).

[16] R. S. Judson, K. K. Lehmann, H. Rabitz, W. S. Warren, *J. Mol. Struct.* **223**, 425 (1990).

[17] T. Schulte-Herbrüggen *Aspects and Prospects of High-Resolution NMR* (Ph.d. Thesis, ETH Zurich, 1998).

[18] Roger W. Brockett, "Explicitly Solvable Control Problems With Nonholonomic Constraints", *IEEE Conference on Decision and Control*, (1999).

[19] B. Kostant, "On Convexity, the Weyl Group and the Iwasawa decomposition" *Ann. Sci. Ecole Norm. Sup.* **6**, 413-455 (1973).